Commodity Machine Learning and beyond

Andreas Müller
NYU, scikit-learn

✉ amueller@nyu.edu
@amuellerml
@amueller
Amueller.io

We Want You!
To Apply Machine Learning!

What ML can do for you

Classification

Hi Andy,

I just received an email from the first tutorial speaker, presenting right before you, saying he's ill and won't be able to make it.

I know you have already committed yourself to two presentations, but is there anyway you could increase your tutorial time slot, maybe just offer time to try out what you've taught? Otherwise I have to do some kind of modern dance interpretation of Python in data :-)
-Leah

Hi Andreas,

I am very interested in your Machine Learning background. I work for X Recruiting who have been engaged by Z, a worldwide leading supplier of Y. We are expanding the core engineering team and we are looking for really passionate engineers who want to create their own story and help millions of people.

Can we find a time for a call to chat for a few minutes about this?

Thanks

Classification

Hi Andy,

I just received an email fr… the first tutorial speaker, presenting … u, saying he's ill and won't be…

I know you have a… …ourself to two presentations… …y you could increase you… …, maybe just offer time to try out what you've taught? Otherwise I have to do some kind of modern dance interpretation of Python in data :-)
-Leah

Hi Andreas,

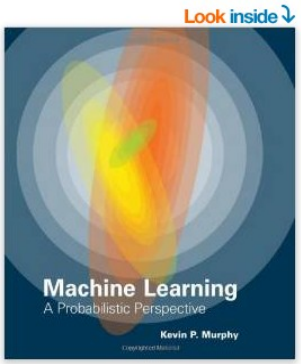
I am very interested in your Machine Learning background. I work for X Recruiting who have been engaged by Z, a worldwide leading supplier of Y. We are expanding the core engineering team and we are looking for really passionate engineers who… …own story and help millions of…

Can we find a ti… …for a few minutes about t…

Thanks

Classification

Recommendations



See all 1 image(s)

Machine Learning: A Probabilistic Perspective (Adaptive Computation and Machine Learning series) Hardcover

by Kevin P. Murphy (Author)

★★★★☆ 31 customer reviews

See all 2 formats and editions

Kindle	Hardcover
$64.06	$77.41

Read with our free app

20 Used from $67.99
42 New from $69.66

Today's Web-enabled deluge of electronic data calls for automated methods of data analysis. Machine learning provides these, developing methods that can automatically detect patterns in data and then use the uncovered patterns to predict future data. This textbook offers a comprehensive and self-contained introduction to the field of machine learning, based on a unified, probabilistic approach. The coverage combines breadth and depth, offering necessary background material on such topics as probability, optimization, and linear algebra as well as discussion of recent developments in the field, including conditional random fields, L1 regularization, and deep learning. The book is written in an informal, accessible style, complete with pseudo-code for the most important algorithms. All topics are copiously illustrated with color images and worked examples drawn from such application domains as

Read more

Share

Buy New $77.41

Qty: 1

List Price: $90.00
Save: $12.59 (14%)

FREE Shipping

In Stock.

Ships from and sold by Amazon.com. Gift-wrap available.

Add to Cart

Sign in to turn on 1-click ordering

Want it Wednesday, June 11? Order within **11 hrs 27 mins** and choose **AmazonGlobal Priority Shipping** at checkout. Details

Buy Used $71.98

Add to Wish List

Trade in your item
Get a **$35.10**
Gift Card.

Trade in

Learn More

Have one to sell?

Sell on Amazon

Frequently Bought Together

Price for all three: $249.90

Add all three to Cart | Add all three to Wish List

Show availability and shipping details

- **This item:** Machine Learning: A Probabilistic Perspective (Adaptive Computation and Machine Learning series) by Kevin P. Murphy Hardcover $77.41
- Pattern Recognition and Machine Learning (Information Science and Statistics) by Christopher M. Bishop Hardcover $85.46
- Probabilistic Graphical Models: Principles and Techniques (Adaptive Computation and Machine Learning … by Daphne Koller Hardcover $87.03

Customers Who Bought This Item Also Bought

Page 1 of 12

Learning From Data
› Yaser S. Abu-Mostafa

Pattern Recognition and Machine Learning …

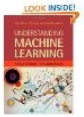
Probabilistic Graphical Models: Principles …

The Elements of Statistical Learning: …

Foundations of Machine Learning (Adaptive …

Machine Learning: The Art and Science of …

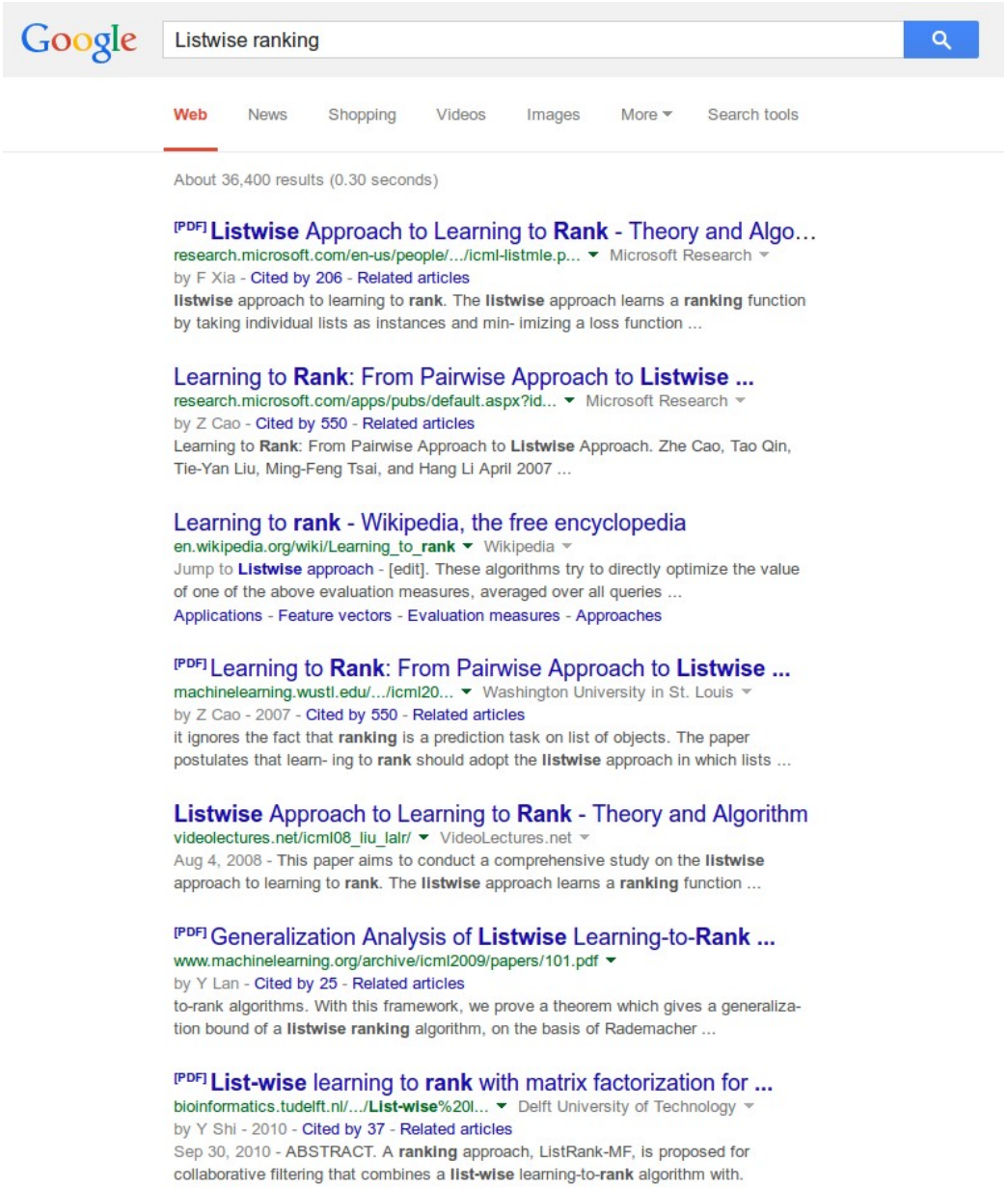
Building Machine Learning Systems with …



Understanding Machine Learning: From Theory to

An Introduction to Statistical Learning: …

Ranking


Google
Listwise ranking
Web
News
Shopping
Videos
Images
More
Search tools
About 36,400 results (0.30 seconds)
[PDF] Listwise Approach to Learning to Rank - Theory and Algo…
research.microsoft.com/en-us/people/.../icml-listmle.p... Microsoft Research
by F Xia - Cited by 206 - Related articles
listwise approach to learning to rank. The listwise approach learns a ranking function by taking individual lists as instances and min- imizing a loss function ...
Learning to Rank: From Pairwise Approach to Listwise ...
research.microsoft.com/apps/pubs/default.aspx?id... Microsoft Research
by Z Cao - Cited by 550 - Related articles
Learning to Rank: From Pairwise Approach to Listwise Approach. Zhe Cao, Tao Qin, Tie-Yan Liu, Ming-Feng Tsai, and Hang Li April 2007 ...
Learning to rank - Wikipedia, the free encyclopedia
en.wikipedia.org/wiki/Learning_to_rank Wikipedia
Jump to Listwise approach - [edit]. These algorithms try to directly optimize the value of one of the above evaluation measures, averaged over all queries ...
Applications - Feature vectors - Evaluation measures - Approaches
[PDF] Learning to Rank: From Pairwise Approach to Listwise ...
machinelearning.wustl.edu/.../icml20... Washington University in St. Louis
by Z Cao - 2007 - Cited by 550 - Related articles
it ignores the fact that ranking is a prediction task on list of objects. The paper postulates that learn- ing to rank should adopt the listwise approach in which lists ...
Listwise Approach to Learning to Rank - Theory and Algorithm
videolectures.net/icml08_liu_lalr/ VideoLectures.net
Aug 4, 2008 - This paper aims to conduct a comprehensive study on the listwise approach to learning to rank. The listwise approach learns a ranking function ...
[PDF] Generalization Analysis of Listwise Learning-to-Rank ...
www.machinelearning.org/archive/icml2009/papers/101.pdf
by Y Lan - Cited by 25 - Related articles
to-rank algorithms. With this framework, we prove a theorem which gives a generaliza- tion bound of a listwise ranking algorithm, on the basis of Rademacher ...
[PDF] List-wise learning to rank with matrix factorization for ...
bioinformatics.tudelft.nl/.../List-wise%20l... Delft University of Technology
by Y Shi - 2010 - Cited by 37 - Related articles
Sep 30, 2010 - ABSTRACT. A ranking approach, ListRank-MF, is proposed for collaborative filtering that combines a list-wise learning-to-rank algorithm with.

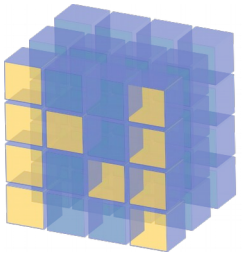
Applying machine learning is easy.

Applying machine learning is easy.
But it should be easier!

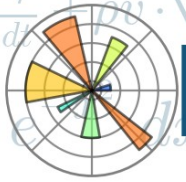
python

IP[y]: IPython
Interactive Computing

SciPy



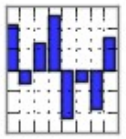
NumPy

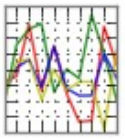


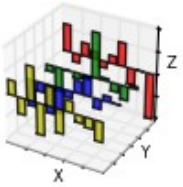
matplotlib

pandas

$$y_{it} = \beta' x_{it} + \mu_i + \epsilon_{it}$$







IP[y]: IPython
Interactive Computing

Chapter 4 - Working With Text Data. - Mozilla Firefox

File Edit View History Bookmarks Tools Help

IPy Chapter 4 - Workin...

127.0.0.1:8888/030923ca-af46-4fc6-86fc-1f0b18130e27#

DuckDuckGo

IP[y]: Notebook Chapter 4 - Working With Text Data. (autosaved)

File Edit View Insert Cell Kernel Help

Code Cell Toolbar: None

In [22]:
```
from sklearn.svm import LinearSVC
svm = LinearSVC(C=.01)
```

In [23]:
```
svm.fit(X_train, y_train)
```

Out[23]:
```
LinearSVC(C=0.01, class_weight=None, dual=True, fit_intercept=True,
     intercept_scaling=1, loss='l2', multi_class='ovr', penalty='l2',
     random_state=None, tol=0.0001, verbose=0)
```

In [24]:
```
svm.score(X_train, y_train)
```

Out[24]:
```
0.88421586014694709
```

In [25]:
```
svm.score(X_test, y_test)
```

Out[25]:
```
0.83679637325273892
```

In [26]:
```
y_test_pred = svm.predict(X_test)
```

In [27]:
```
from sklearn.metrics import classification_report
```

In [28]:
```
print(classification_report(y_test, y_test_pred))
```

```
             precision    recall  f1-score   support

          0       0.84      0.96      0.90      1954
          1       0.82      0.48      0.61       693

avg / total       0.83      0.84      0.82      2647
```

In [29]:
```
coef = svm.coef_.ravel()
positive_coefficients = np.argsort(coef)[-25:]
negative_coefficients = np.argsort(coef)[:25]
interesting_coefficients = np.hstack([negative_coefficients, positive_coefficients])
```

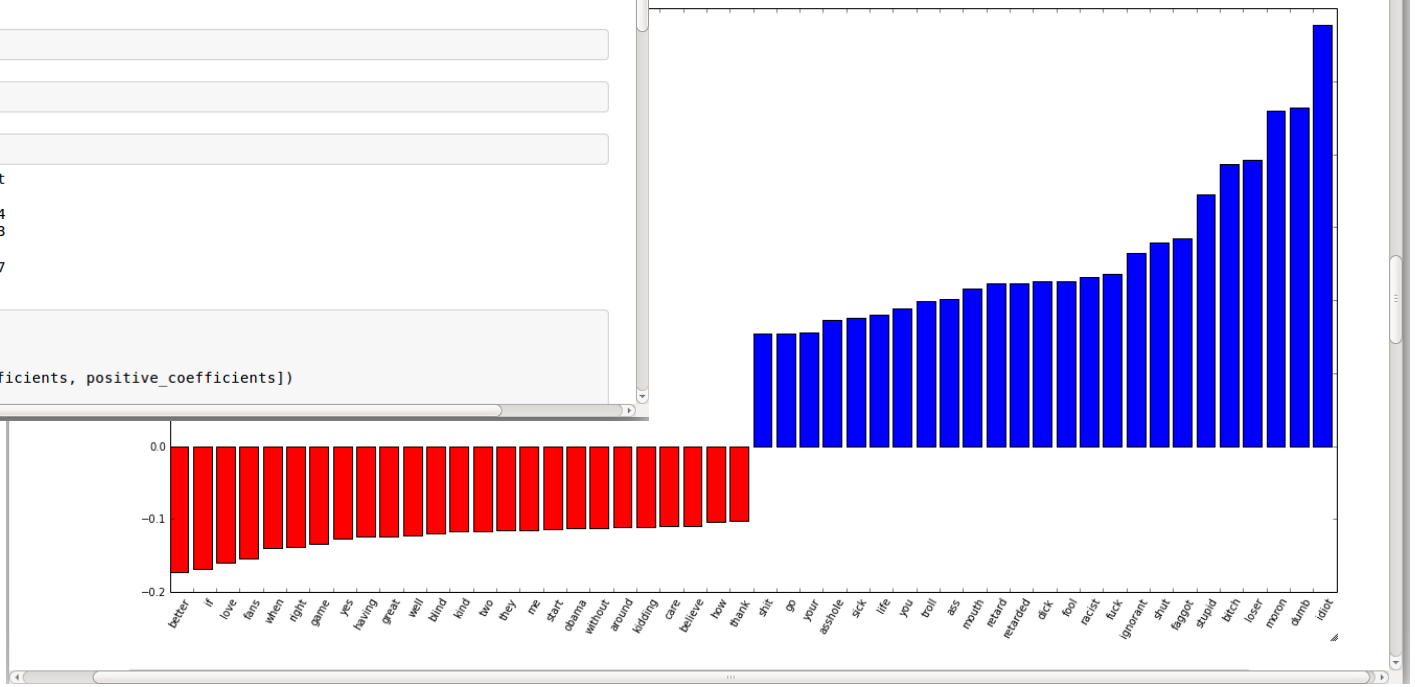
DuckDuckGo

xt Data. (autosaved)

ents, positive_coefficients])

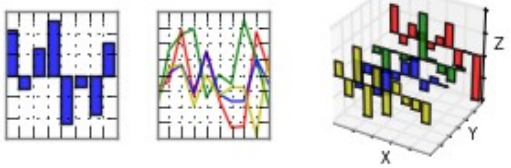
color=["red" if c < 0 else "blue" for c in coef[interesting_coefficients]])

coefficients], rotation=60, ha="right");

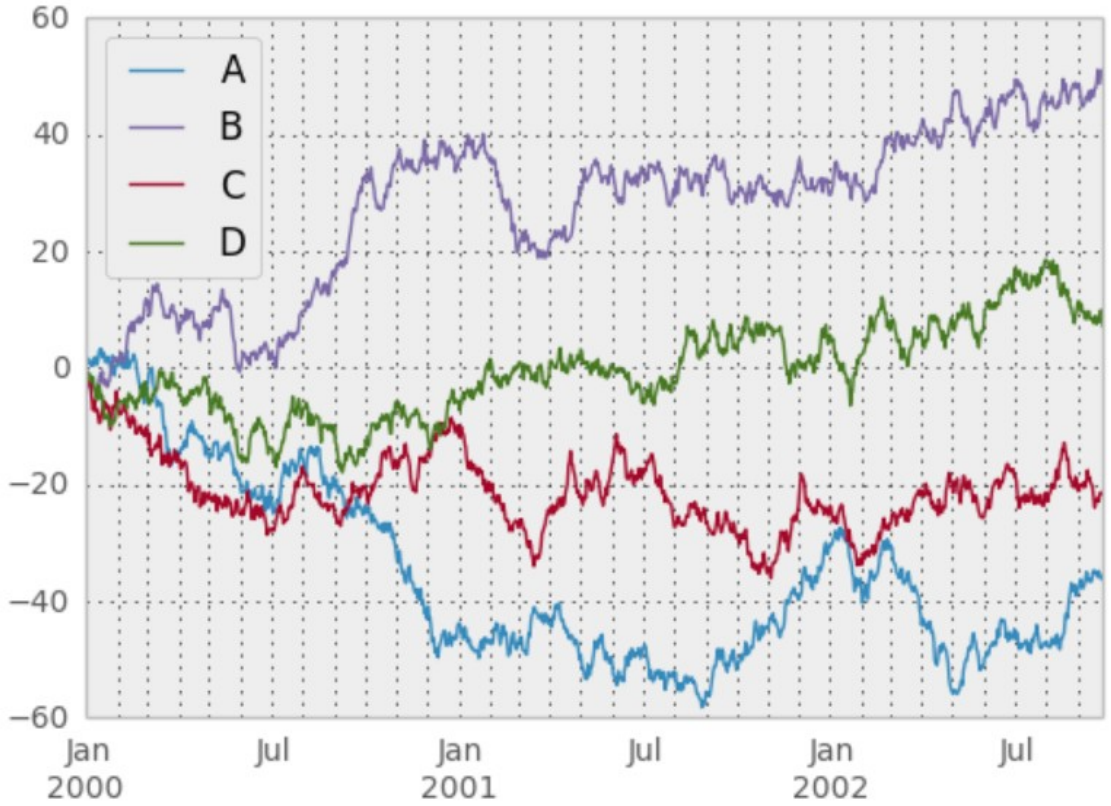


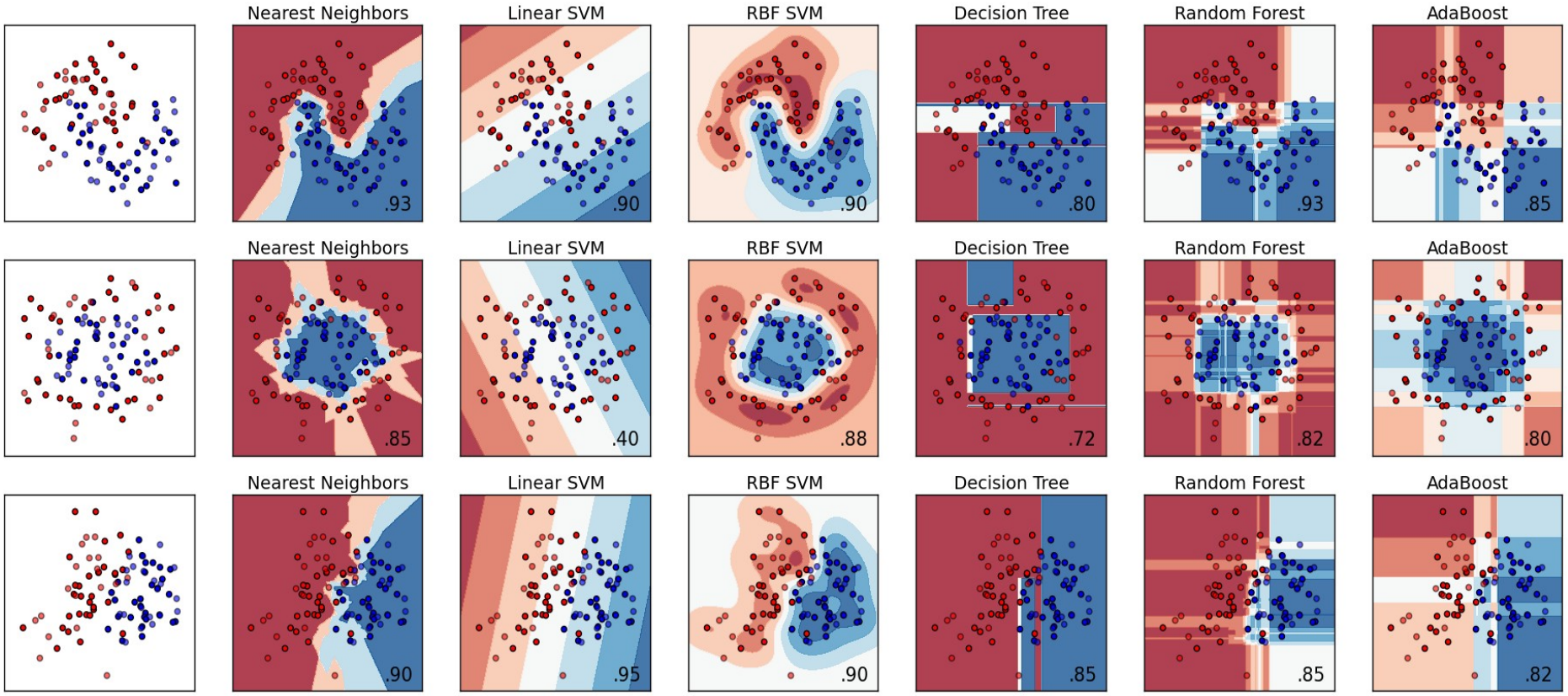
pandas

$$y_{it} = \beta' x_{it} + \mu_i + \epsilon_{it}$$




```
In [6]: df = DataFrame(randn(1000, 4), index=ts.index, columns=list('ABCD'))

In [7]: df = df.cumsum()

In [8]: plt.figure(); df.plot();
```




Nearest Neighbors
Linear SVM
RBF SVM
Decision Tree
Random Forest
AdaBoost
.93
.90
.90
.80
.93
.85
Nearest Neighbors
Linear SVM
RBF SVM
Decision Tree
Random Forest
AdaBoost
.85
.40
.88
.72
.82
.80
Nearest Neighbors
Linear SVM
RBF SVM
Decision Tree
Random Forest
AdaBoost
.90
.95
.90
.85
.85
.82

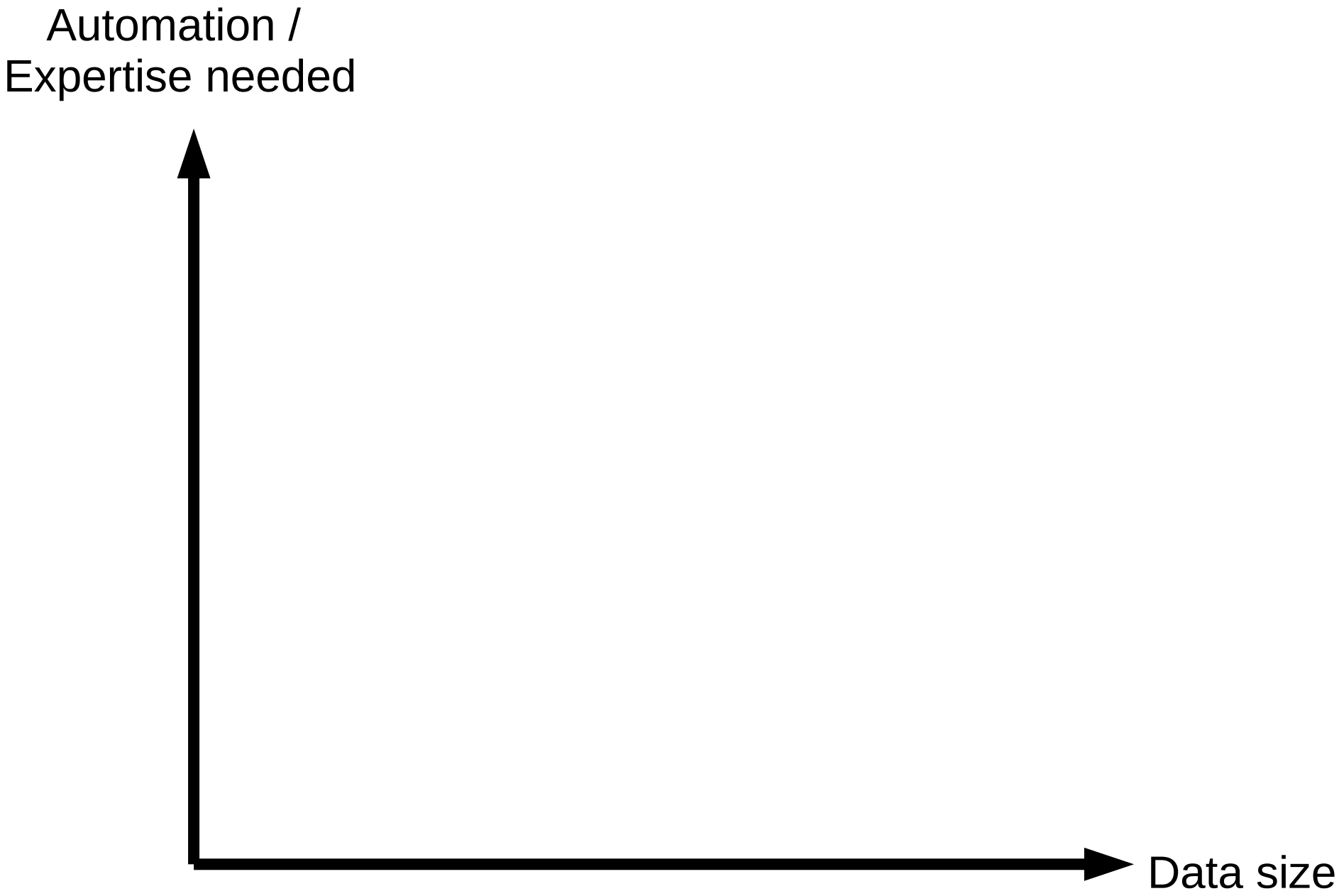
2750+ research papers

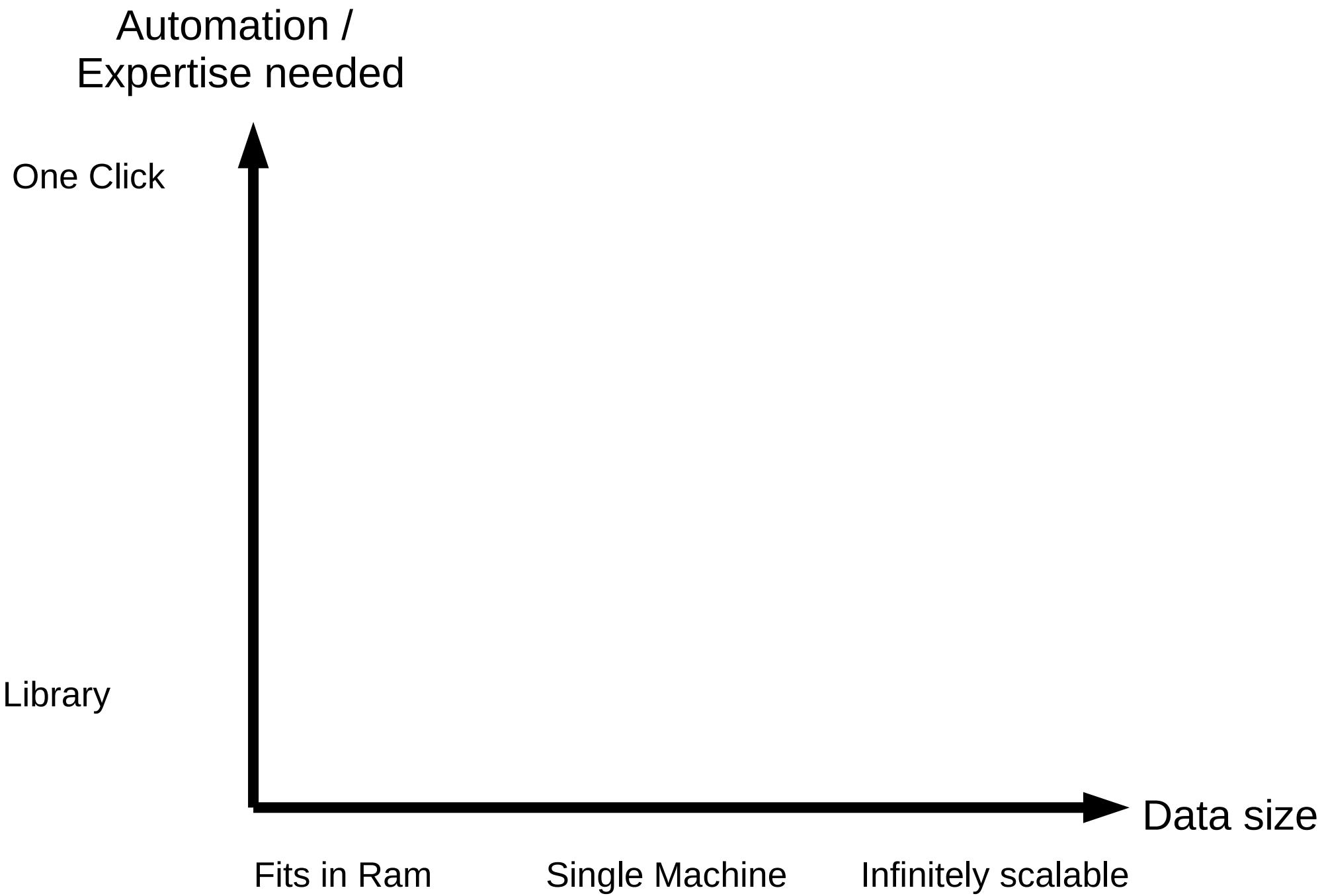
Fully Functional Spam Classifier

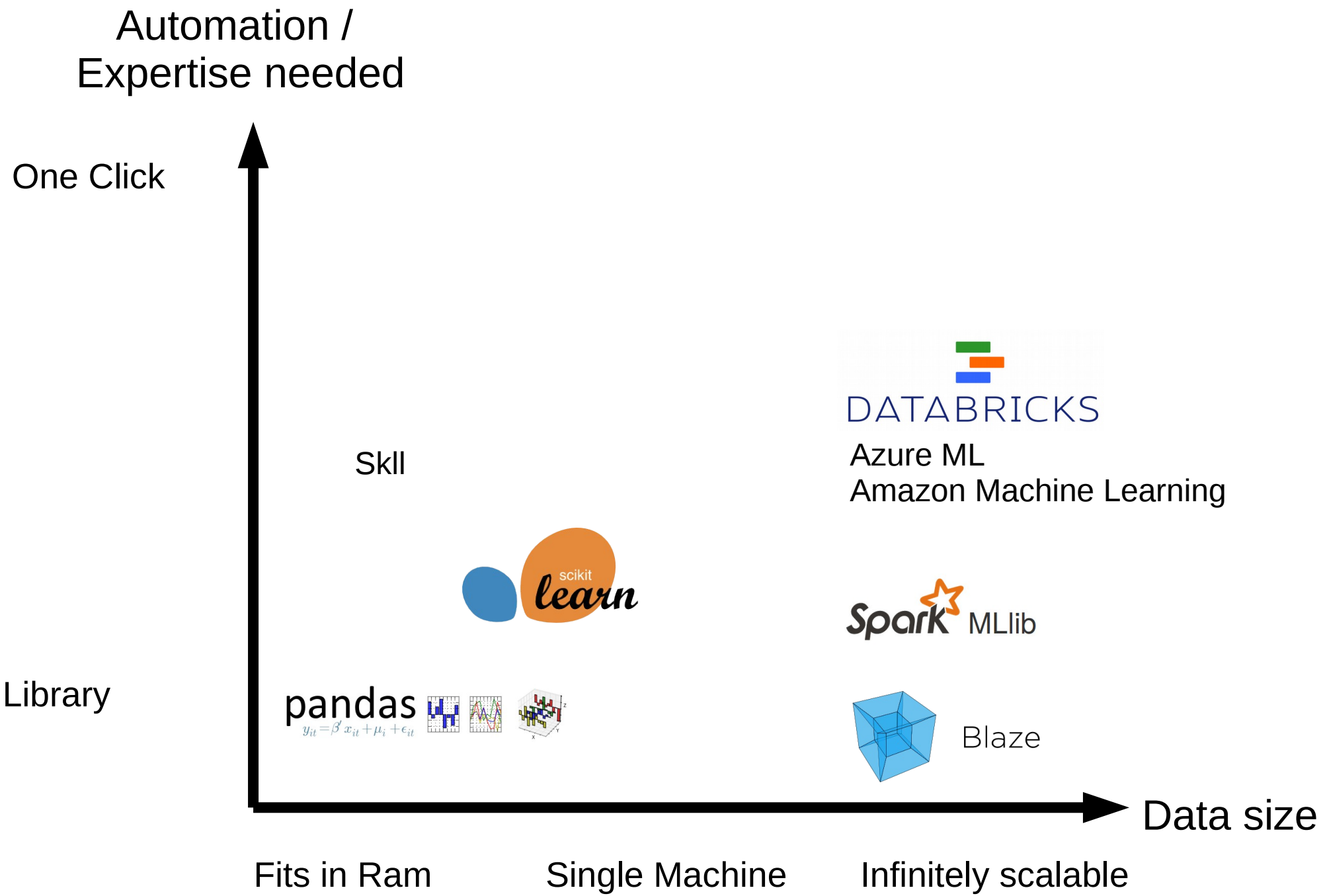
```
from sklearn.naive_bayes import MultinomialNB
from sklearn.feature_extraction.text import CountVectorizer
from pipeline import make_pipeline

spam_classifier = make_pipeline(CountVectorizer(),
                                MultinomialNB())
spam_classifier.fit(email_texts, is_spam)

spam_classifier.predict(new_emails)
```

“The scikit-learn tutorials / documentation is so good, one doesn't need a textbook anymore to learn a new machine learning method.”

This is not enough!


Automation /
Expertise needed
Data size


Automation /
Expertise needed
One Click
Library
Data size
Fits in Ram
Single Machine
Infinitely scalable


Automation /
Expertise needed
One Click
Library
SkIl
DATABRICKS
Azure ML
Amazon Machine Learning
scikit
learn
Spark MLlib
pandas
Blaze
Data size
Fits in Ram
Single Machine
Infinitely scalable

A single machine is (usually) enough.

	vCPU	ECU	Memory (GiB)	Instance Storage (GB)	Linux/UNIX Usage
r3.large	2	6.5	15	1 x 32 SSD	$0.166 per Hour
r3.xlarge	4	13	30.5	1 x 80 SSD	$0.333 per Hour
r3.2xlarge	8	26	61	1 x 160 SSD	$0.665 per Hour
r3.4xlarge	16	52	122	1 x 320 SSD	$1.33 per Hour
r3.8xlarge	32	104	244	2 x 320 SSD	$2.66 per Hour
Storage Optimized - Current Generation					
i2.xlarge	4	14	30.5	1 x 800 SSD	$0.853 per Hour
i2.2xlarge	8	27	61	2 x 800 SSD	$1.705 per Hour
i2.4xlarge	16	53	122	4 x 800 SSD	$3.41 per Hour
i2.8xlarge	32	104	244	8 x 800 SSD	$6.82 per Hour
d2.xlarge	4	14	30.5	3 x 2000 HDD	$0.69 per Hour
d2.2xlarge	8	28	61	6 x 2000 HDD	$1.38 per Hour
d2.4xlarge	16	56	122	12 x 2000 HDD	$2.76 per Hour
d2.8xlarge	36	116	244	24 x 2000 HDD	$5.52 per Hour

We need open box methods.

We need black-box methods.

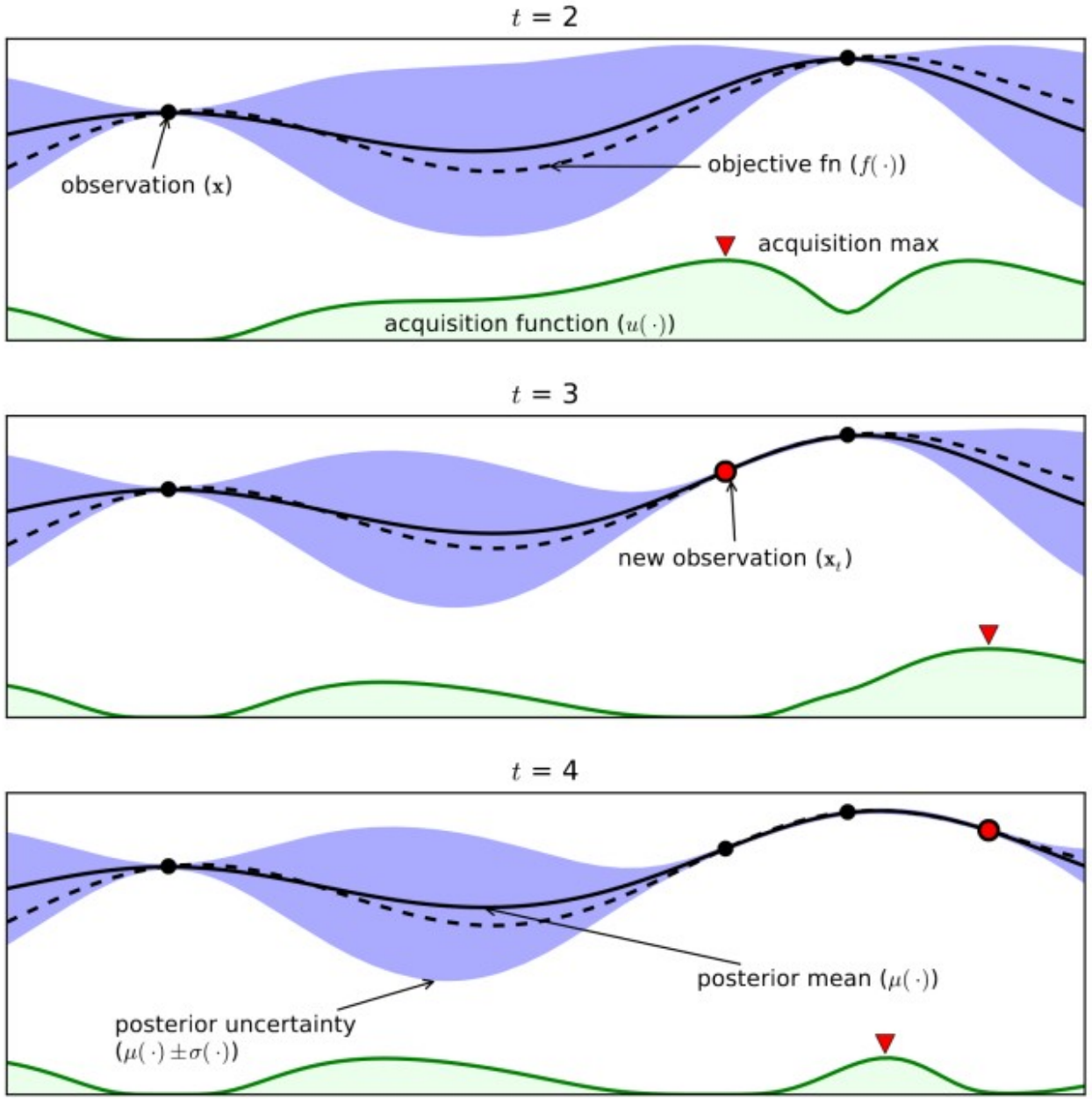
predict

Hyperparameter Optimization

Spearmint
Hyperopt
SMAC

… scikit-learn

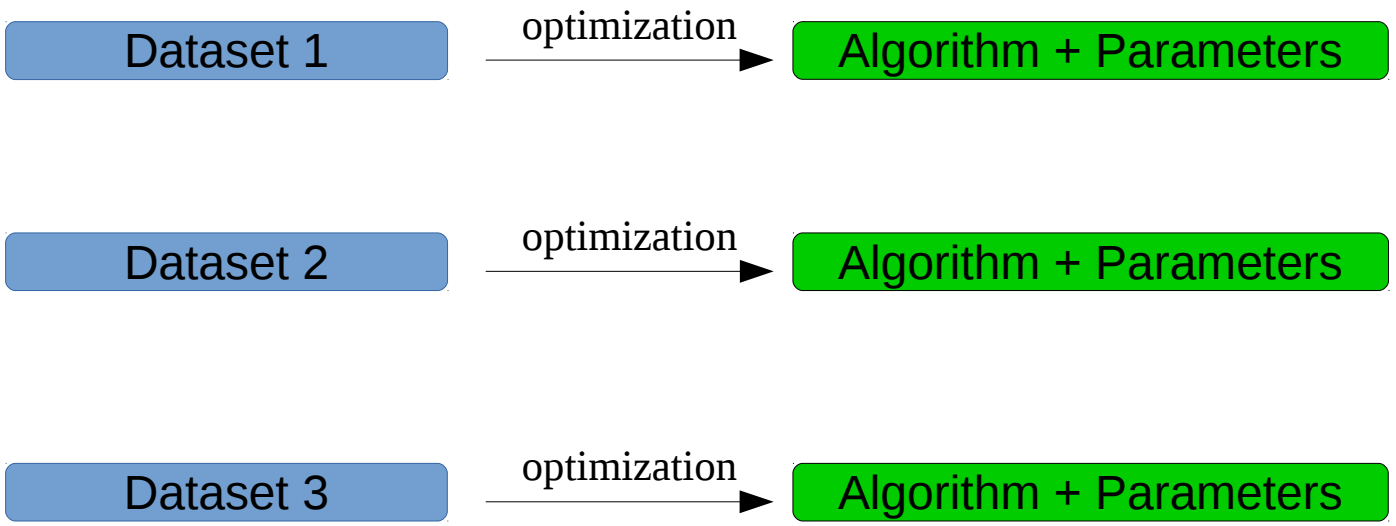
Bayesian Optimization



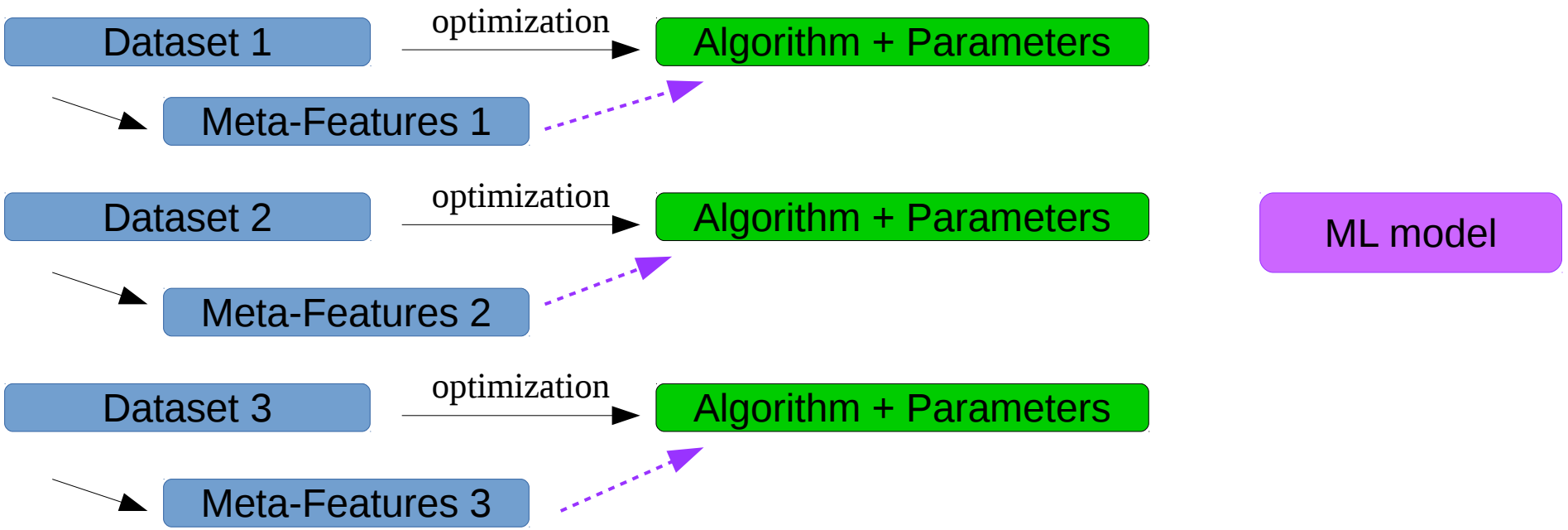
From Eric Brochu, Vlad M. Cora and Nando de Freitas

Meta-Learning

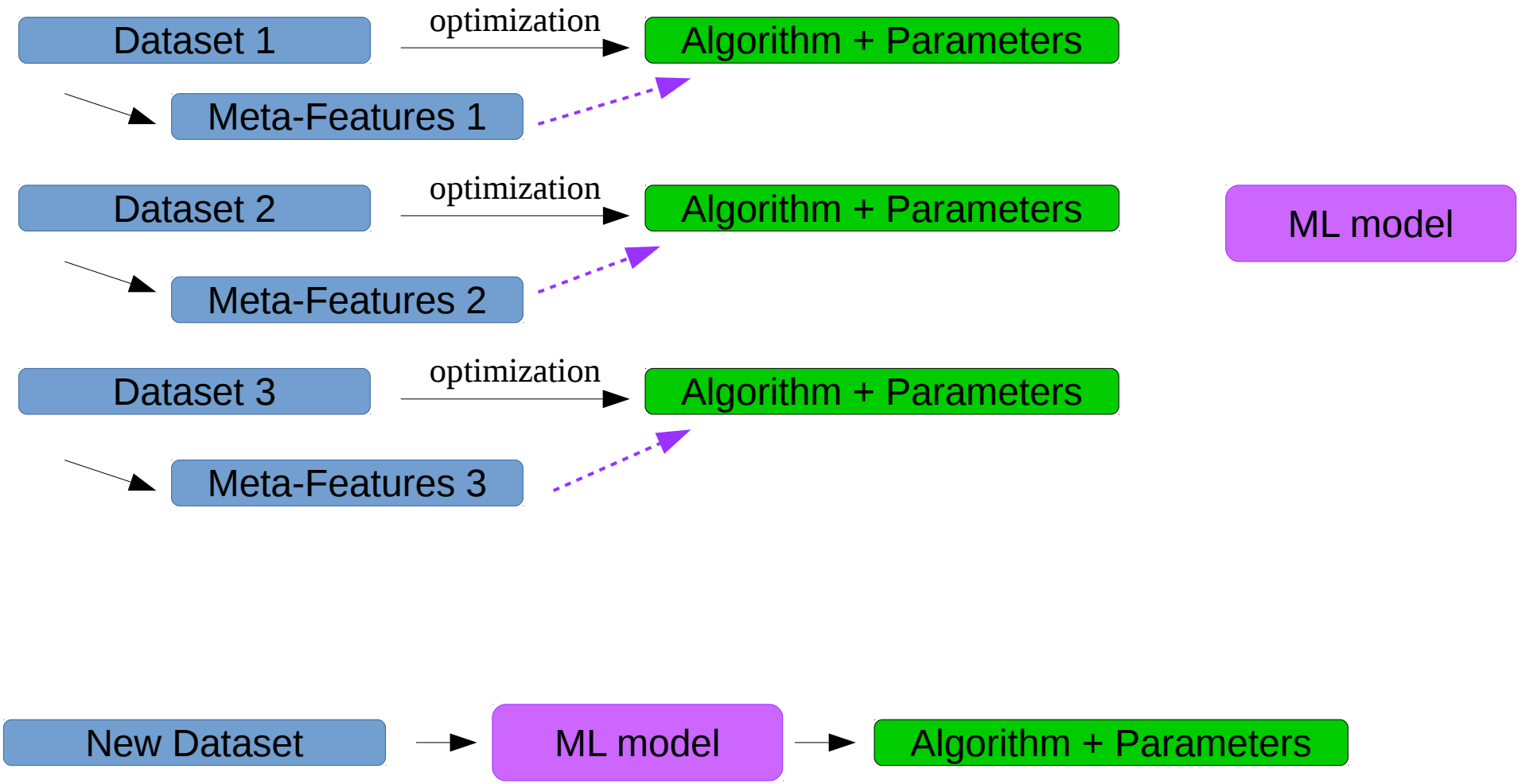
Meta-Learning

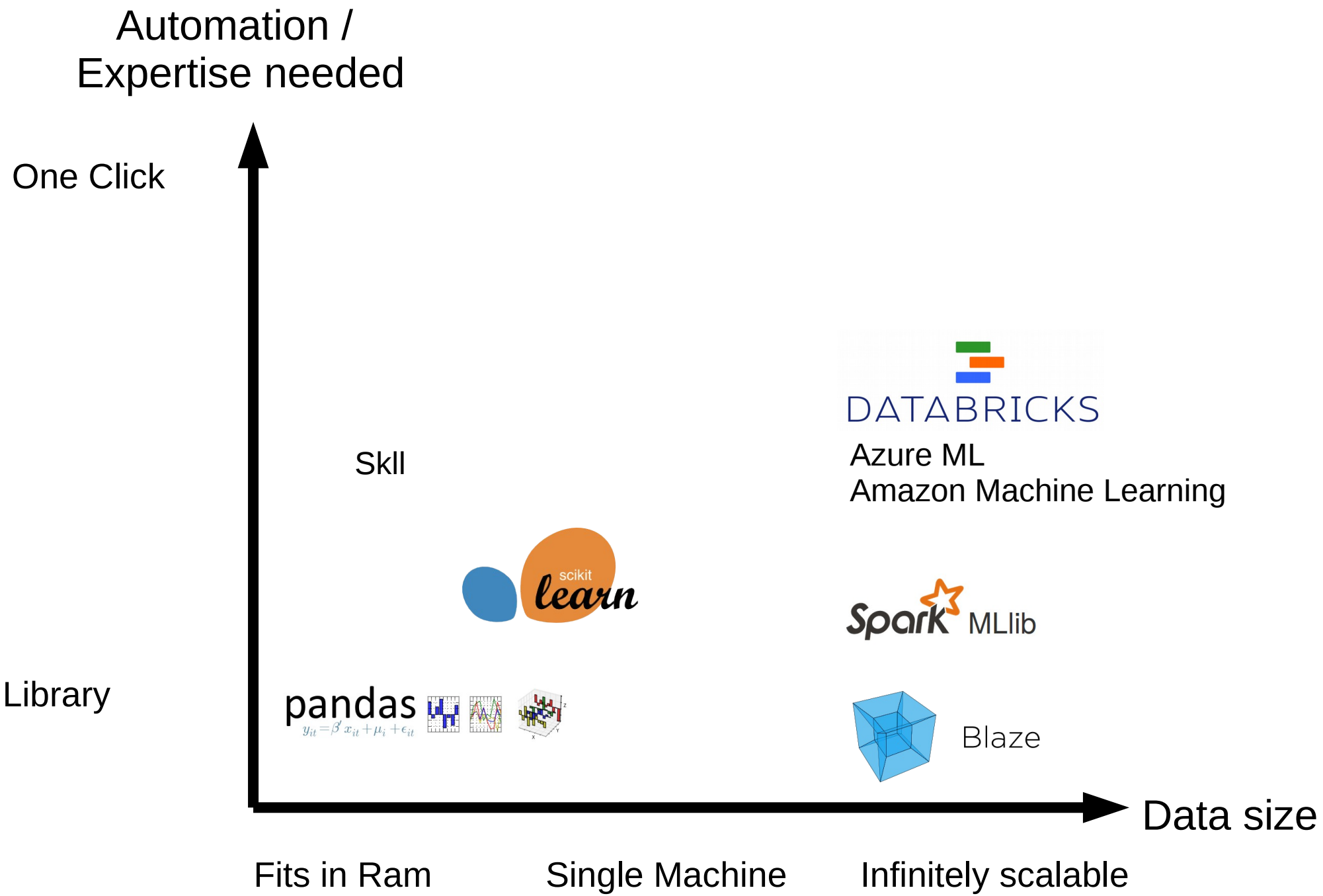


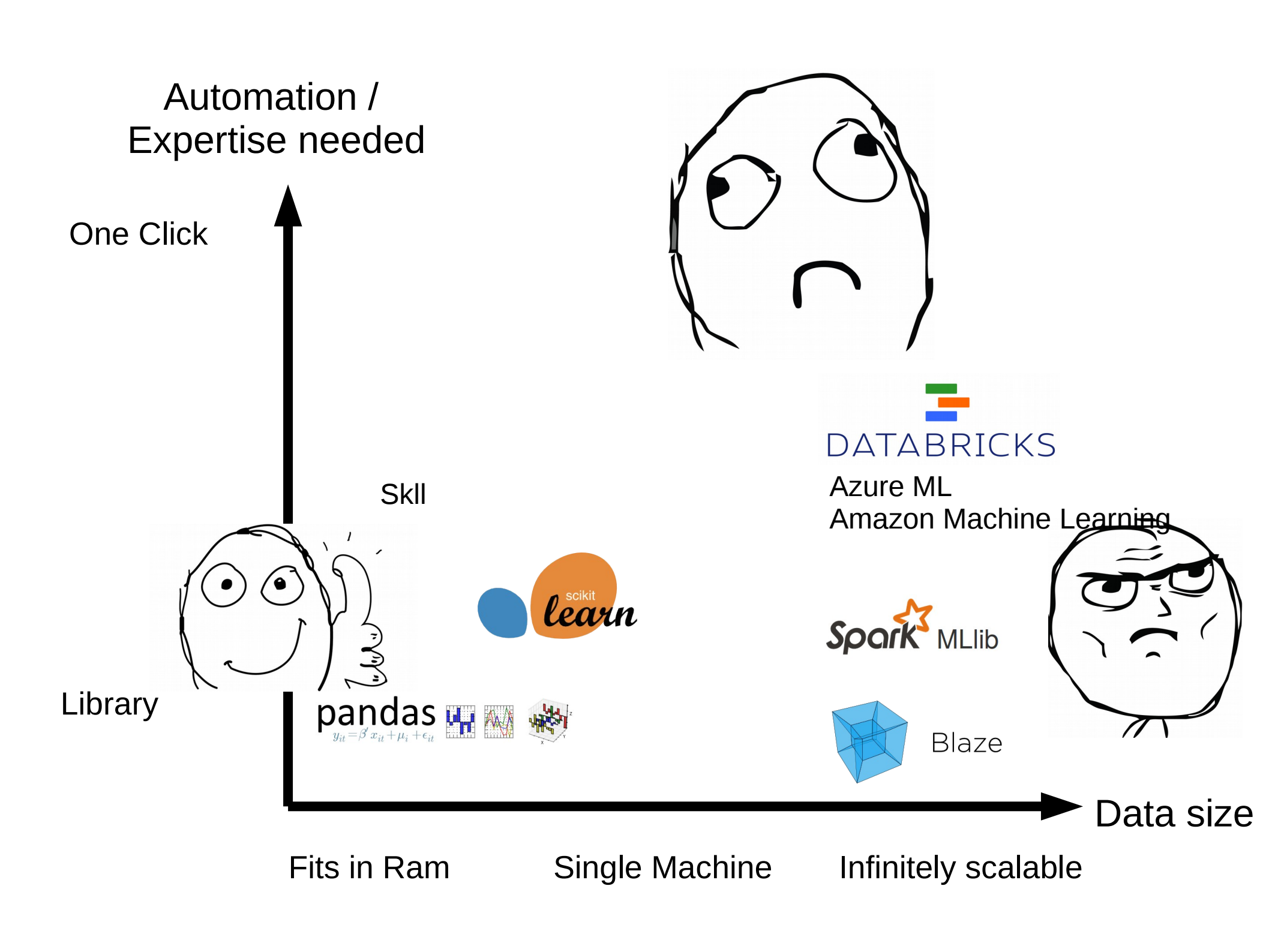
Meta-Learning

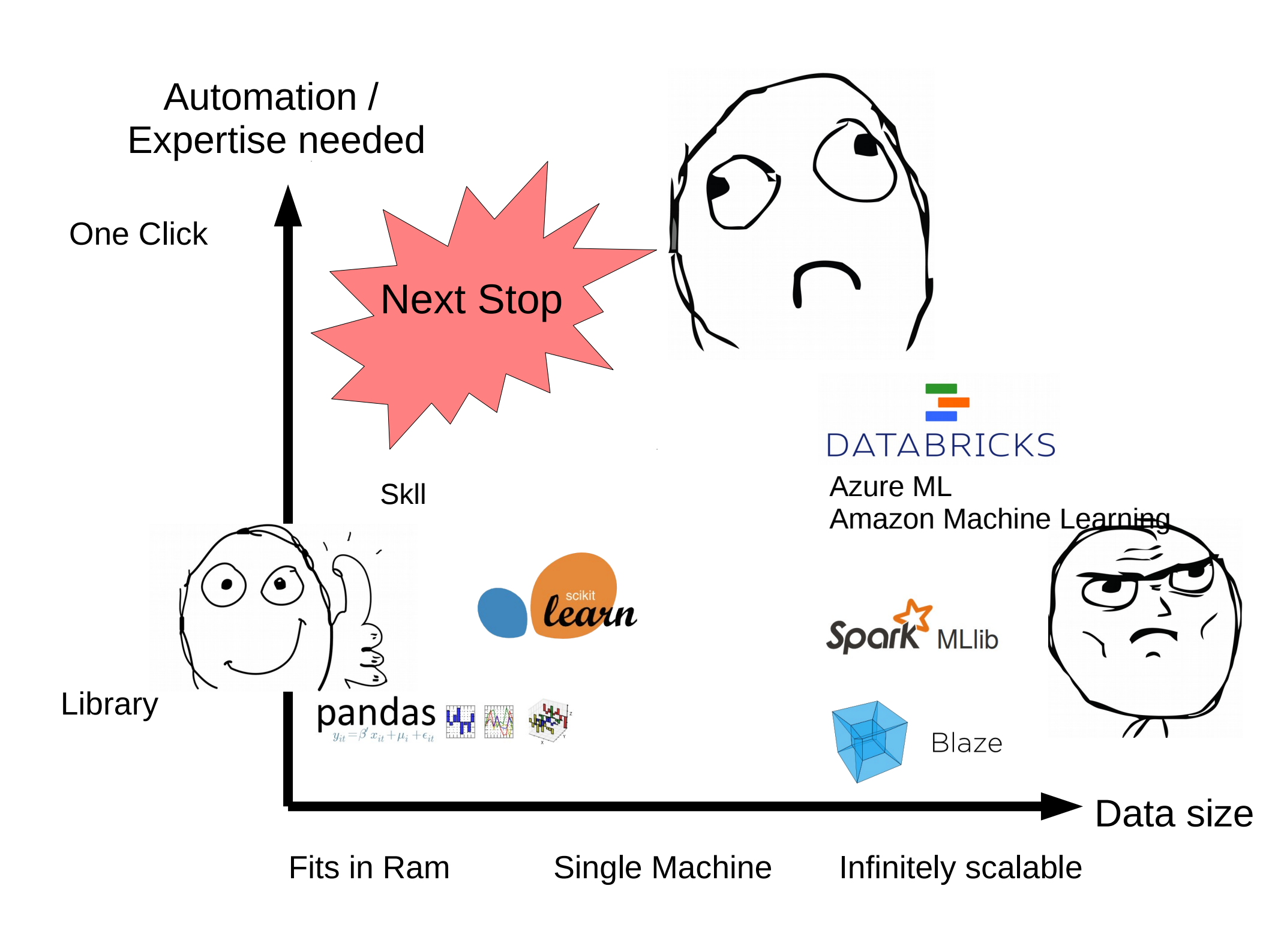


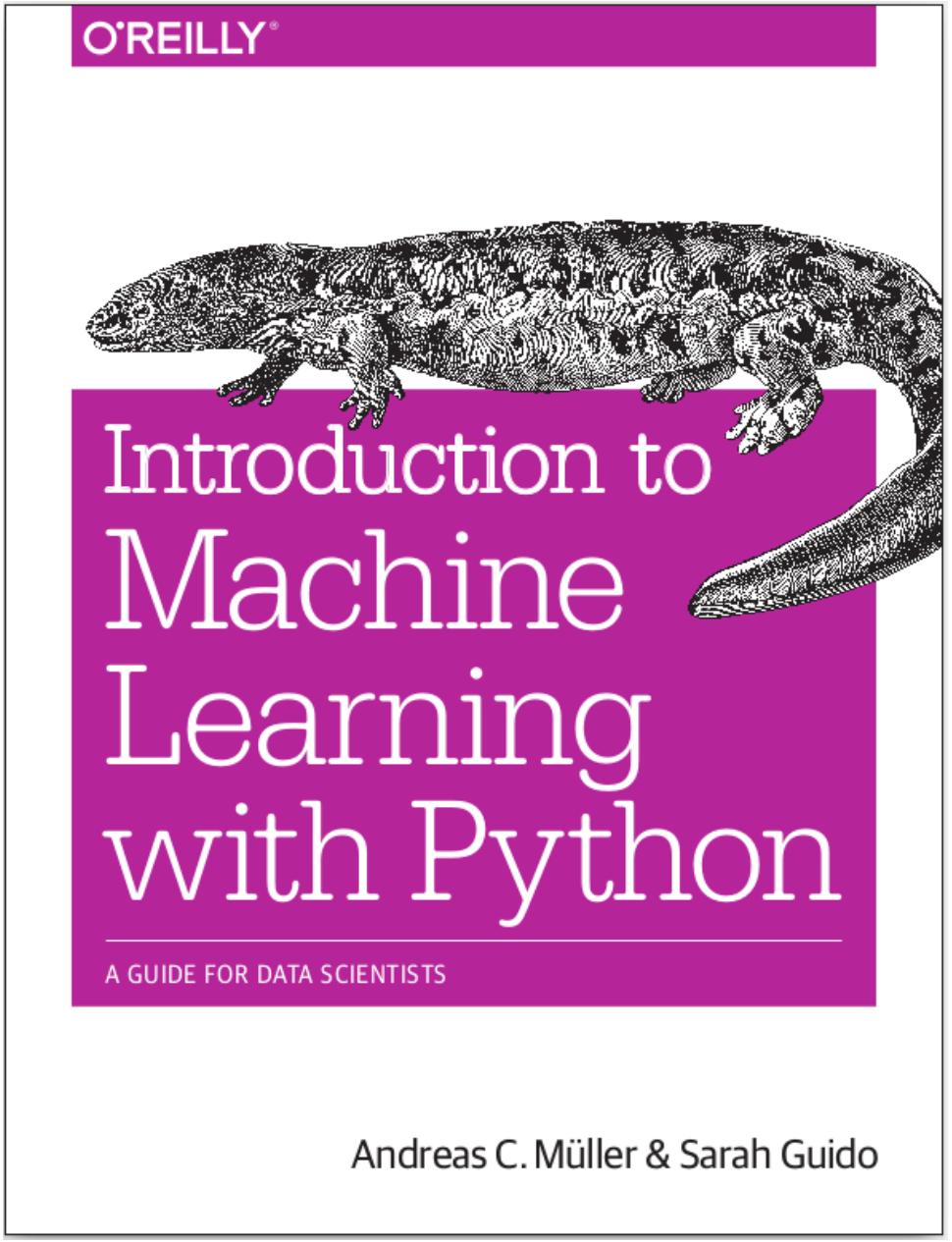
Meta-Learning




Automation /
Expertise needed
One Click
Library
SkIl
DATABRICKS
Azure ML
Amazon Machine Learning
scikit
learn
Spark MLlib
pandas
Blaze
Data size
Fits in Ram
Single Machine
Infinitely scalable


Automation /
Expertise needed
One Click
Skll
Library
DATABRICKS
Azure ML
Amazon Machine Learning
scikit
learn
pandas
Spark MLlib
Blaze
Data size
Fits in Ram
Single Machine
Infinitely scalable


Automation /
Expertise needed
One Click
Next Stop
Skll
Library
DATABRICKS
Azure ML
Amazon Machine Learning
scikit
learn
Spark MLlib
pandas
Blaze
Data size
Fits in Ram
Single Machine
Infinitely scalable



Release June 2016

Thank you.

@amuellerml

@amueller

amueller@nyu.edu

http://amueller.io